## CLAIMS

What is claimed:

1      1.     A method for organizing and aiding the interpretation of gene data, said method
2 comprising the steps of:
3      receiving gene names;
4      associating the gene names to gene-word pair relationships; and
5      grouping the gene names with high strength of gene-word relationships, the strength of
6 the gene-word relationships corresponding to the relatedness in function of corresponding
7 grouped genes.

1      2.     The method of claim 1, wherein the receiving gene names includes receiving alias
2 names for the gene names.

1      3.     The method of claim 1, further including querying the gene names in a literature
2 database.

1      4.     The method of claim 3, wherein the receiving includes, responsive to the query of
2 the gene names in a literature database, receiving abstracts comprising the gene names.

1      5.     The method of claim 4, further including generating a background set and a query
2 set from the returned abstracts.

1      6.     The method of claim 5, further including calculating word frequencies in the
2 query set and the background set.

1      7.     The method of claim 6, further including providing a numerical value calculated
2 for each word in which a word frequency was calculated for the query set.

1     8.     The method of claim 7, wherein the providing includes calculating z scores.

1     9.     The method of claim 7, wherein the providing includes using term frequency-
2     inverse document frequency methods.

1     10.     The method of claim 4, further including stemming words of the returned
2     abstracts.

1     11.     The method of claim 10, further including filtering the stemmed words using a
2     stop list.

1     12.     A system for organizing and aiding the interpretation of data, said system
2     comprising:
3            a memory with logic; and
4            a processor configured with the logic to receiving gene names, said processor further
5     configured with the logic to associate the gene names to gene-word pair relationships, said
6     processor further configured with the logic to group the gene names with high strength of gene-
7     word relationships, the strength of the gene-word relationships corresponding to the relatedness
8     in function of corresponding grouped genes.

1     13.     The system of claim 12, wherein the processor is further configured with the logic
2     to generate keywords that describe the common function of each group.

1        14.     A system for organizing and aiding the interpretation of gene data, said system

2    comprising:

3          means for receiving gene names;

4          means for associating the gene names to gene-word pair relationships; and

5    means for grouping the gene names with a similar strength of gene-word relationships, the

6    strength of the gene-word relationships corresponding to the relatedness in function of

7    corresponding grouped genes.